AFRL-ML-WP-TP-2007-430

PROGNOSTIC FUSION FOR UNCERTAINTY REDUCTION (Preprint)

Kai Goebel, Neil Eklund, and Pierino Bonanni

FEBRUARY 2007

Approved for public release; distribution unlimited.

STINFO COPY

This work was funded in whole or in part by Department of the Air Force work unit M02R3000. The U.S. Government has for itself and others acting on its behalf an unlimited, paid-up, nonexclusive, irrevocable worldwide license to use, modify, reproduce, release, perform, display, or disclose the work by or on behalf of the U. S. Government.

MATERIALS AND MANUFACTURING DIRECTORATE
AIR FORCE RESEARCH LABORATORY
AIR FORCE MATERIEL COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433-7750

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YY)*
February 2007

2. REPORT TYPE
Conference Preprint

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
PROGNOSTIC FUSION FOR UNCERTAINTY REDUCTION (Preprint)

5a. CONTRACT NUMBER
IN HOUSE

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
62102F

6. AUTHOR(S)
Kai Goebel, Neil Eklund, and Pierino Bonanni

5d. PROJECT NUMBER
4347

5e. TASK NUMBER
RG

5f. WORK UNIT NUMBER
M02R3000

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
GE Global Research
One Research Circle
Niskayuna, NY 12309

8. PERFORMING ORGANIZATION REPORT NUMBER
AFRL-ML-WP-TP-2007-430

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Materials and Manufacturing Directorate
Air Force Research Laboratory
Air Force Materiel Command
Wright-Patterson AFB, OH 45433-7750

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
AFRL-ML-WP

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-ML-WP-TP-2007-430

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
PAO Case Number: AFRL/WS 07-0235, 06 February 2007. Paper contains color. Submitted for publication in the proceedings of the AIAA Infotech at Aerospace 2007 Conference and Exhibit, May 10, 2007, Rohnert Park, CA. This work was funded in whole or in part by Department of the Air Force work unit M02R3000.

14. ABSTRACT
This paper describes how the fusion of two different prognostic approaches produces a result that is more accurate and has more narrow uncertainty bounds than either approach alone. The fused prognostic estimate can be calculated by using both a physics-based as well as a data-driven approach. The individual approaches can have a plurality of input sources such as component properties (e.g., material properties and usage properties), history of the component (current damage state and history of accumulated usage), future anticipated usage, damage propagation rates established during experiments, etc. Damage estimates are arrived at using sensor information such as oil debris monitoring data as well as vibration data. The method detects the onset of damage and triggers the prognostic estimator that projects the remaining life. Uncertainty, stemming from the variability observed during experiments, as well as modeling inaccuracies, are propagated to provide a distribution around the projected remaining life. It is desirable to keep the uncertainty interval as narrow as possible while truthfully considering their spread. In this paper, we introduce an approach to fuse competing prediction algorithms for prognostics. Results presented are derived from rig test data wherein multiple bearings were first seeded with small defects, then exposed to a variety of speed and load conditions similar to those encountered in aircraft engines, and run until the ensuing material liberation accumulated to a predetermined damage threshold or cage failure, whichever occurred first.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor) / 19b. TELEPHONE NUMBER *(Include Area Code)*
Unclassified	Unclassified	Unclassified	SAR	16	Andrew H. Rosenberger (937) 255-1305

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

Prognostic Fusion for Uncertainty Reduction

Kai Goebel, Neil Eklund, Pierino Bonanni
GE Global Research
One Research Circle
Niskayuna, NY 12309
518-387-4194
{ goebelk, eklund, bonanni} @ research.ge.com

Abstract**—**This paper describes how the fusion of two different prognostic approaches produces a result that is more accurate and has more narrow uncertainty bounds than either approach alone. The fused prognostic estimate can be calculated by using both a physics-based as well as a data-driven approach. The individual approaches can have a plurality of input sources such as component properties (e.g., material properties and usage properties), history of the component (current damage state and history of accumulated usage), future anticipated usage, damage propagation rates established during experiments, etc. Damage estimates are arrived at using sensor information such as oil debris monitoring data as well as vibration data. The method detects the onset of damage and triggers the prognostic estimator that projects the remaining life. Uncertainty, stemming from the variability observed during experiments, as well as modeling inaccuracies, are propagated to provide a distribution around the projected remaining life. It is desirable to keep the uncertainty interval as narrow as possible while truthfully considering their spread. In this paper, we introduce an approach to fuse competing prediction algorithms for prognostics. Results presented are derived from rig test data wherein multiple bearings were first seeded with small defects, then exposed to a variety of speed and load conditions similar to those encountered in aircraft engines, and run until the ensuing material liberation accumulated to a predetermined damage threshold or cage failure, whichever occurred first.

TABLE OF CONTENTS

INTRODUCTION

Reasoners attempt to analyze a variety of information sources towards a particular goal. In this case, the goal of the reasoner is to provide a remaining life estimate. To that end, it negotiates and aggregates independent information sources while taking their inherent uncertainty into account. The uncertainty varies as a function of time, the priors on reliability of the information sources, domain knowledge, among others

There are numerous approaches to accomplish aggregation of information such as bagging and boosting [Freund and Schapire, 1999], Dempster-Shafer [Smets, 1994], model-based approaches [Nelson and Mason, 1999], fuzzy fusion [Loskiewicz and Uhrig, 1994] or statistics-based approaches [Rao, 2000]. However, it has to be realized that the aggregation itself is only one function of the overall reasoner. In addition to combining information, it has to be ensured that the information that is being used provides the maximum information content. There are a number of issues that need to be dealt with prior to the actual aggregation. Specifically, the information needs to be checked for consistency, and it needs to be cleaned of outliers, noise, faulty or otherwise bad sensor information, it needs to be conditioned and formatted to allow a proper comparison. In addition to that, special cases need to be taken into account that, depending on the situation, should be done either before or after the actual aggregation step. To assist in these tasks, one can employ a sequential and parallel multi-layered configurations strategy. Elements from this configuration strategy have been proven successful in diagnostic fusion environments within project IMATE [Ashby and Scheuren, 2000]. There, a hierarchical, multi-layer architecture [Goebel et al., 2004] was demonstrated that implemented some of these concepts. Information from various diagnostic models and evidential information sources was combined and manipulated through a series of steps that increased and decreased the weight given to the information sources according to the strategies implemented in the respective layers of the fusion process.

An approach more closely related to this paper is non-parametric regression (NPR). Here, no assumptions about the underlying functional form are made. NPR is characterized by low bias (i.e, it can easily represent

underlying function) but at the expense of high variance (i.e., the model will change from realization to realization of the data). That in turn may change the response dramatically depending on data. The simplest idea is the k-nearest neighbor regression that results in good fit, but huge variance and discontinuous behavior. Kernel regression [Watson, 1964; Nadaraya, 1964] overcomes some of these shortcomings by locally weighting members closer to the value in question.

Classical regression techniques (including kernel regression, MLP, RBF, splines, linear, etc.) assume perfect knowledge of *y* (both precise and certain). However, these techniques do not work optimally if knowledge of sensor measurement *y* is imprecise due to limited precision and accuracy of sensors, and if sensor measurement *y* is uncertain (e.g., due to sensor failure). The issue is exacerbated when there are multiple sensors with different sensitivities and reliabilities. In situations where the probe point is very different from that employed in the training set it might be desirable to have mechanisms to cast doubt on the validity of the output.

Dempster Shafer regression [Petit-Renaud and Denoeux, 2004] (DSR) provides a prediction of the output in form of a fuzzy belief assignment. This assignment is defined as a collection of fuzzy sets of values with associated masses of belief. The output is computed using a nonparametric, instance-based approach: evidence samples $e_i = (x_i, m_i)$ in the neighborhood of the input vector *x* are sources of partial information on the response variable. The evidence samples can be represented by a fuzzy belief assignment $m_y[x, e_i]$. Relevance of the evidence with respect to *y* is assumed to be dependent on the dissimilarity to *y*. If *x* is “close” to x_i according to a given metric ||.||, *y* is expected to be close to y_i, which makes example e_i quite relevant to predict the value of *y*. On the contrary, if *x* and x_i are very dissimilar, example e_i provides only marginal information regarding the value of *y*. Therefore, neighborhood evidence input elements are discounted as a function of their distance to *x*. They are then pooled using Dempster’s rule of combination. While the method can cope with heterogeneous training data, the more important characteristics in this context is the formalism for modeling both unreliable and imprecise information provided by multi-sensor systems.

DSR determines the value of sensor measurement *y* at a given time by discounting the belief mass of each observation by:

$$\phi\left(\left|x - x_i\right|\right) = \gamma e^{-\frac{(x-x_i)^2}{\Theta^2}}$$

where:

γ is a tuning parameter (usually >= 0.9)

Θ is a scale parameter, commonly set using cross validation on training data

Next, the discounted belief masses are combined using appropriate version of DS combination. When there are many data points, the computational overhead can become considerable. A remedy is using only the k-nearest neighbors to reduce the complexity of the calculation with little loss of accuracy.

However, Dempster-Shafer regression does not, amongst other things, address how to integrate the future estimated variability of the estimators.

Application to Bearing Damage

During bearings operation, initially localized spalls can initiate that may grow and ultimately result in loss of function. Important factors affecting damage initiation and damage propagation are changes in bearing loads, speeds, and environment. Lubrication, presence of material defects, surface degradation, and external contamination all factor in to the bearing environment. Subsurface fatigue cracks are induced at locations of peak shear stress, become surface-connected, and lead to eventual liberation of material. It is important to assess the microstructural evolution, environmental embrittlement, cyclic hardening, and residual stress to calculate the propagation of bearing damage. The current state is determined by feeding direct sensor data and indirect parameters computed from sensor data into an ensemble of diagnostic algorithms as a basis for input to the fault-evolution and life models [Littles and Buczek, 2004]. The algorithms arrive at their conclusion either by direct measurement, models supported by measurements, or are simply triggered by measurements. The information sources that the reasoner relies on may be updated at different intervals during or between flights and may have different prediction horizons.

Prognostics is about the estimation of remaining useful life under particular assumptions of future use. Sensor measurements provide instantaneous feedback on current damage levels and form the foundation for prognostic estimates. Ideally, features derived from sensor measurements would have monotonically changing properties that accurately reflect increasing component damage and be provided irrespective of external conditions. However, in practice this is nearly never the case: features reflect the noise inherent in sensed data and react differently during particular stages of damage evolution (e.g., some are useful for fault detection, but not for damage growth tracking).

Oil debris monitor features, such as particle counts, have excellent tracking properties that are invariant to changes of environmental parameters [Dempsey et al., 2002]. However, they may be not as suitable to identify fault initiation because their resolution is too low for small damage levels. In addition, absolute counts can be misleading when material gets trapped over time and due to external

contamination. Better sensors for fault initiation and initial fault growth tracking may be vibration sensors that have the promise to pick up smaller damage levels. Features from various transforms such as Fourier, Hilbert, and Wavelets can be useful in detecting and categorizing incipient faults. The vibration sensor's capacity for early detection comes at the price of sensitivity to environmental effects [Dempsey et al., 2002] that are sometimes difficult to quantify or correct. In an aircraft engine, and in particular under conditions of military use, these changes can be significant.

It is thus expedient to aggregate vibration and oil debris information to take advantage of the benefits of both. The fusion of information from oil debris and vibration information, along with knowledge about system and machinery history can result in interactions that may improve the confidence about system condition [Byington et al., 1999]. Howard and Reintjes [Howard and Reintjes, 1999] describe the benefits of using several information sources for fault detection, and discuss oil debris and vibration for helicopter gearboxes in particular. Byington et al. [Byington et al., 1999] describe a fusion technique that correlates the failure mode phenomena with appropriate features. Dempsey et al. [Dempsey et al., 2002] report on the use of fuzzy logic to integrate oil debris and vibration information for gearbox faults where the output was quasi-action recommendations such as "OK, inspect, shutdown".

Prognostic Information Fusion

Different approaches can be employed to estimate future damage. One is to model from first principles the physics of the system as well as the fault propagation for given load and speed conditions. Such a model must include detailed knowledge of material properties, thermodynamic behavior, etc. Alternatively, an empirical experience-based model can be employed wherein data from experiments at known conditions and component damage level are used to build a model for fault propagation rate. Such a model relies heavily on a reasonably large set of experiments that sufficiently explores the load and speed space.

The two approaches for estimating future damage have various advantages and disadvantages. The physics-based model relies on the assumption that the fault mode modeled using the specific geometry, material properties, temperature, load, and speed conditions will be similar to the actual fault mode. Deviation in any of those parameters will likely result in an error that is amplified over time. In contrast, the experience-based model assumes that the data available sufficiently maps the space and that interpolations (and extrapolations) from that map can capture the fault rate properly. It can be beneficial to fuse the output of both methods to produce a more robust and more accurate result. Finding synergy in using different information sources to assess system states has a long tradition within the fields of multivariate statistics and pattern recognition.

In addition to fusing a damage estimate, the associated uncertainty needs to be aggregated as well. This is a critical task because the resulting estimate needs to be within uncertainty bounds that allow for decision making at a desired risk level. If the uncertainty bounds are very wide, the resulting time-of-failure estimate at the acceptable risk level may be too early to provide any benefit to the decisioning process. That is, there would be no advantage of prognostics compared to a reactionary diagnostics system alone. Uncertainty bounds ideally are tight but need to reflect true output variability.

Prognostic Fusion Techniques

The aggregation of future damage estimates is not just a question of averaging the various values. Rather, the fusion method should be able to incorporate a number of different measures that inform about the reliability of the estimate, their expected accuracy, and various other uncertainty measures. These measures in turn may be a function of different variables such as time, where in the load/speed space the estimate is performed, known shortcomings or strength in some areas of that space, etc. In the example described by Orsagh et al. [Orsagh et al., 2003], performance improvement is accomplished when weights for the information sources are dynamically allocated depending on whether the component is considered early or late in its remaining useful life cycle. Garga et al. [Garga et al., 2001] describe a hybrid reasoning approach that integrates domain knowledge with test and operational data from an industrial gearbox. There, domain knowledge is expressed as a rule-base, and then used to train a feedforward neural network.

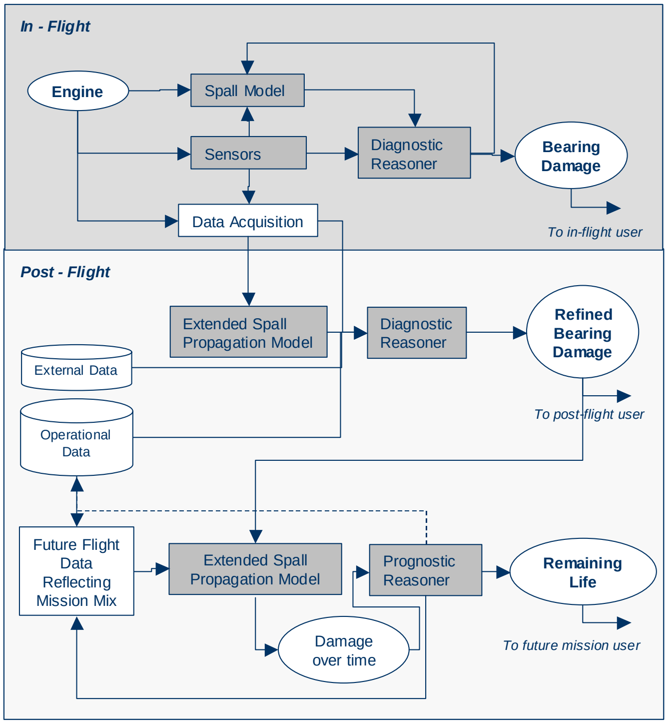


Figure 1 - Interactions of Integrated Bearing Reasoner Modules

ARCHITECTURE

As mentioned above, the prognostic reasoner considered here is really a set of reasoners that will operate at various times during and after the flight. Depending on the time during or after a mission, its tasks will vary from aggregation of damage information to supporting the calculation of a remaining life estimate. There are two fundamentally different modes: one is a diagnostic mode that estimates the magnitude of the fault. Another is the prognostic mode that establishes a time horizon for remaining life. The two modes are described in more detail below:

<u>In-Flight and Post-Flight Diagnostic Modes:</u> Using an in-mission setting, features derived from vibration measurements and debris counts are used in transfer functions to provide a damage detection indicator. Specifically, an adaptive neuro-fuzzy inference scheme (ANFIS) was used that takes these information sources as input and gives fault presence likelihood *fp* as output:

$$fp = f\left(features_{debris}, features_{vibration}\right)$$

where

f = neuro fuzzy inference system.

ANFIS is a technique invented by Roger Jang in 1993 [Jang, 1993]. Any other suitable mapping function can be employed here as well, such as neural nets, support vector machines, random forests, etc. The detection algorithm is tuned to avoid false positives and to minimize late detection. If the output of the fault presence exceeds a fault detection threshold, the fault is declared present.

Next (and only after the fault has been detected) a suite of transfer functions converts sensor-based features into equivalent damages $\hat{d}_{debris}, \hat{d}_{vibration}$, for debris-based damage estimates and vibration-based damage estimates, respectively. Again ANFIS or other suitable mapping function can be employed. Specifically, we used ANFIS here:

$$\hat{d}_i = f\left(features_i\right)$$

where

i is either the debris information or the vibration information.

Additional damage estimates come from an experience-based tool (described in more detail below) as well as a physics-based tool. In parallel, quality estimates are provided for each estimate. The quality estimate is a subjective assessment for the goodness of the output.

The diagnostic functions are displayed in the flowchart in Figure 2.

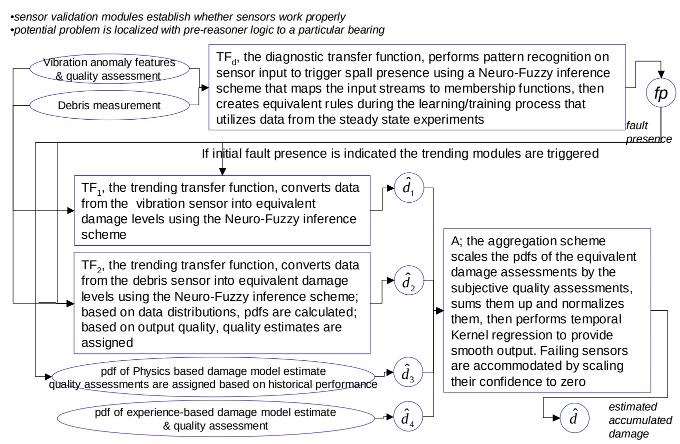


Figure 2 – Diagnostic Flowchart

Next, an aggregator combines the information, trading off the quality estimates and fusing the pdf-based information. The fusion is the focus of this paper and will be described in more detail below.

<u>Prognostic Mode</u> The prognostic models can be run either on-board or on-ground, depending on whether there is a need for short-term outlook (in which case the prognostic reasoner would be executed on-board) or whether there is a need for a longer-term outlook (in which case it makes more sense to run the prognostic reasoner on-ground). If a fault has been detected, the prognostic functions are executed on

a set of future missions. Specifically, missions characterized in part by sequences of load, speed, and ambient conditions are used as input to the physics-based spall propagation model as well as the experience-based model. In conjunction with the current damage state, the output of the spall propagation model will provide a damage profile into the future.

The modeled damage over time and the quality assessment over time from each model are then forwarded to the aggregation module. Figure 3 illustrates the operation of the prognostic reasoner. Fundamentally, the prognostic reasoner supervises the execution of the different prognostic models, makes corrections where desired, and assigns a quality assessment. It then aggregates the different estimates. There are different ways in which the reasoner can operate based on user demand. In one instantiation, it will report both the profile of remaining life and information on whether the envisioned missions can be completed without exceeding the acceptable damage limit. In another instantiation, it will provide information back to the mission generation process to prompt for additional mission runs when damage limits have not been reached. The goal of executing the damage propagation model with additional runs is to determine the damage propagation profile and to find the remaining life limit.

As mentioned before, if no fault has been detected, the prognostic module is bypassed and is replaced by fleet statistics that are compiled on bearing fatigue data.

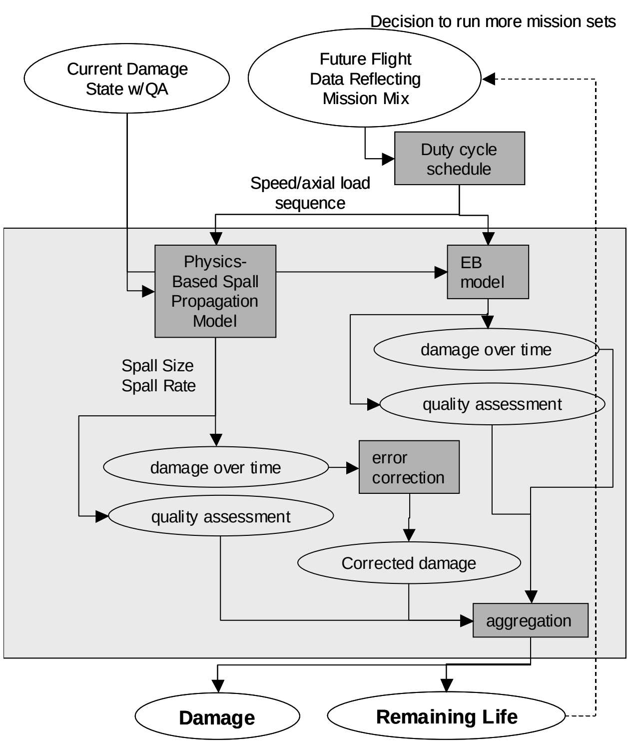


Figure 3 - Prognostic Reasoner

METHOD

Below we will provide a detailed description of the method. We discuss preprocessing, assignment of quality estimates, estimation of variability, the experience-based prognostic model, aggregation of uncertainty, and postprocessing.

Assignment of quality estimates

In addition to the damage estimate, each model is assigned a quality assessment that can be interpreted as a subjective confidence. These confidences are computed based on *a priori* performance of the models. That is, the models may be known to have a different performance within different regions of the load-speed mission space. Additionally, the models may be known to produce biases at different damage levels or at different damage rate levels. Moreover, the further out into the future the prediction is being made, the less likely it is to be correct. While confidence intervals may capture the possible variability, the quality assessment captures other sources of uncertainty. If one takes into account the *quality* of the model (e.g., derived by examining performance of the model) for particular regions of the search space (or other factors, e.g., time), one has the possibility to exploit this additional information during the aggregation step which ultimately may result in better performance of the prognostics. This was discussed in detail in [Goebel et al., 2006].

Experience-based Prognostic Model

Two models are fused in the prognostic reasoner, a physics-based (PB) model and an experience-based (EB) model.

The EB model is an empirical fit of data from seven experiments at five points in the speed and load space. Spall length is calculated:

$$l_{spall} = 10^{\log10(l_{spall_{t=0}}) + \sum_{t=0:dt:current} rate(t)*dt}$$

where

$$rate = 10^{f(speed(t), load(t))}.$$

Spall growth rate is exponential, with rate an empirical function of speed and load. Spall rate was calculated from the raw data, and a surface was fit using a relatively simple (to avoid unwanted distortions in the surface) neural network (two input nodes, two hyperbolic tangent hidden nodes, and one linear output node). Figure 4 is a plot of the response of the model to individual test runs. Figure 5 is a plot of the response surface of the model showing the data it was modeled from; Figure 6 is another view of the response surface.

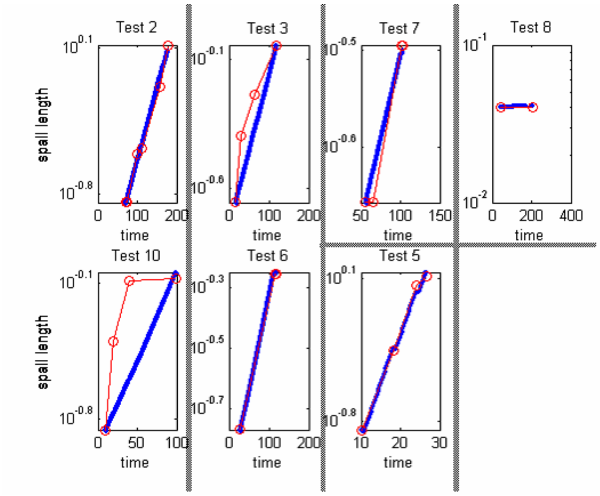


Figure 4 - Response of the model to individual test runs. Red is actual data, blue is model predicted spall length. Grey lines join tests with the same conditions.

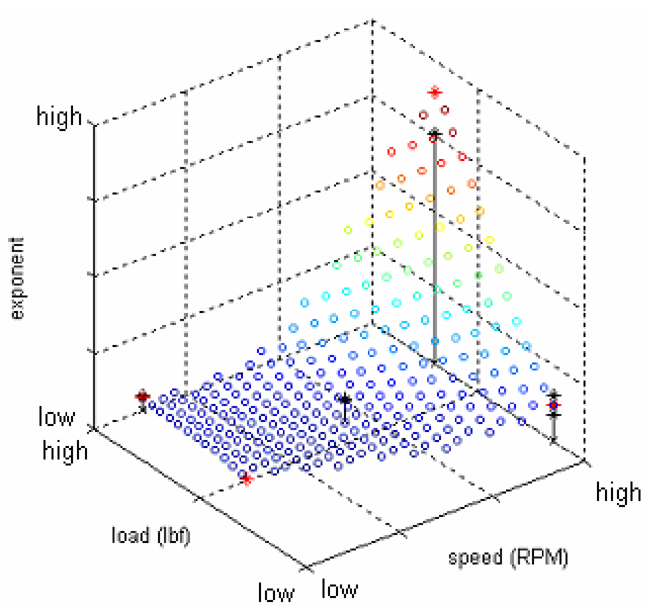


Figure 5 - Response surface of experience-based model showing raw data.

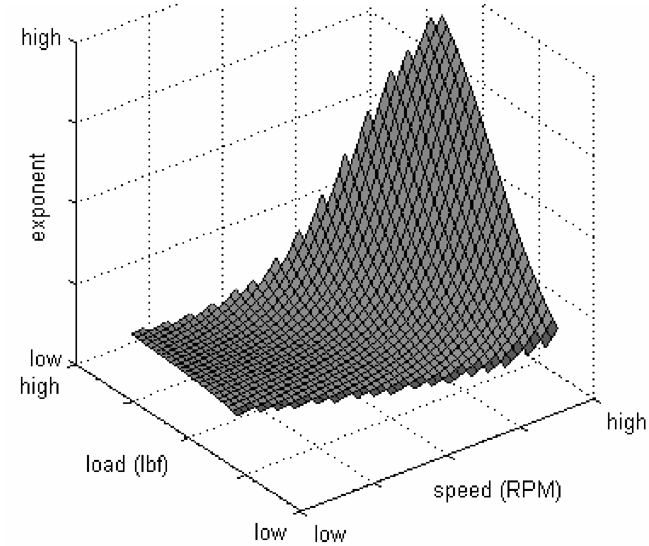


Figure 6 - Response surface of experience-based model.

Physics-Based Prognostics Model

The PB model for the initiation and propagation of bearing fatigue spall uses historic and estimated future operating conditions to determine future bearing condition and returns a probability density function of the bearing remaining useful life. This model is based on first principles approaches such as damage mechanics to track material microstructure changes and eventual loss during the spall propagation phase. It takes into account material properties, bearing geometry, surface interaction, lubrication, and variable operating conditions.

The physics-based model was developed by Sentient Corp [Marble et al., 2006]. We added on an error correction to the model at the time of prognostics. Due to the open-loop calculation of the PB model, the damage estimate at the time of prognostics may have an offset compared to the best damage of the reasoner. This may lead to a propagation of that bias throughout the prognostic horizon. To counteract that, the reasoner subtracts the bias of the PB-based mean estimate at the time of prognostics from the reasoner-based mean estimate at the time of prognostics.

Aggregation

The primary goal of the prognostic reasoner is to negotiate the different damage estimates and to decide whether another set of mission parameters needs to be executed for another damage estimate further in the future. A key to the reasoner's performance is the ability to aggregate different measures of uncertainty. To this end, we propose a new method as described in the following.

To properly aggregate multiple estimates of spall size, it is necessary to account for both model uncertainty and model quality assessment, as discussed above, and to

accommodate model updates at arbitrary (possibly different or asynchronous) updating intervals.

All spall length estimates are put on a common time scale using interpolation, which accommodates different or asynchronous model updating times. Each estimate PDF is then discretized at each time interval over a finely divided (e.g., 1000 intervals) universe of discourse (at most 0% to 100% of race length, but often much less, depending on the maximum non-zero values of all spall length estimate PDFs). The discretized PDF of each estimate is discounted by its unique time-dependent quality assessment values.

$$discounted\_pdf_t = qa_t * pdf_t$$

The discounted PDFs are aggregated using kernel regression (i.e., discounting events distant in time from the time currently being evaluated) using

$$pdf_{aggregated} = \frac{\sum_{i=1}^{N} K_\lambda(t_0, t_i) \cdot discounted\_pdf_i}{\sum_{i=1}^{N} K_\lambda(t_0, t_i)}$$

where

$$K_\lambda = \begin{cases} \left(1 - \frac{|t_i - t_0|}{\lambda}\right)^2 & \text{for } \frac{|t_i - t_0|}{\lambda} < 1 \\ 0 & \text{otherwise} \end{cases}$$

Finally, the aggregate PDF is renormalized at each time interval, and the desired spall length percentiles are returned. The basic concept is illustrated in Figure 7:

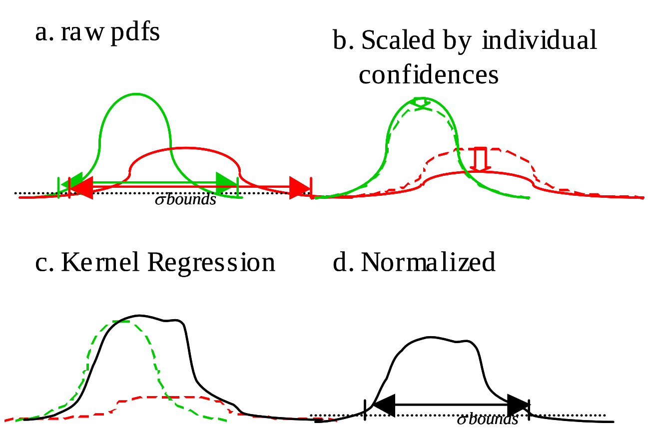


Figure 7 – Aggregation Concept

First, the raw probability density functions (Figure 7a) are scaled by the individual quality estimates (Figure 7b). Next, the PDFs are combined using kernel regression (Figure 7c) and normalized (Figure 7d). The resulting spread of the fused PDF is smaller than the original ones at the same level of risk (say, 3 σ) as illustrated in Figure 8.

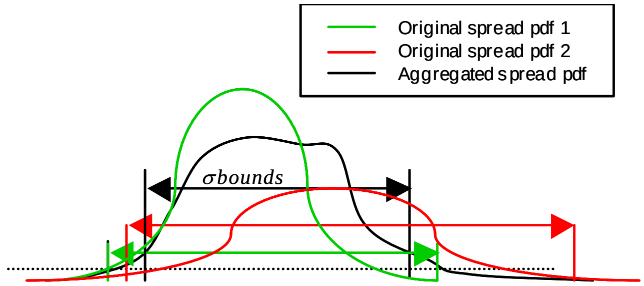


Figure 8 – Spread of original pdf's and aggregated pdf

Postprocessing

Some output of the damage estimate transfer functions can be noisy. That in turn may result in suboptimal behavior in the fusion function. Specifically, it is undesirable to have non-monotonic behavior. To reduce noise and encourage monotonic properties, an adaptive filter was employed that is responsive to increases while being more cautious to downward changes of the input. Specifically, an exponential weighted moving average filter was employed where weight α was modified based on the situation at hand. The governing equation is:

$$damage_{debris}(k) = \alpha \cdot damge_{debris_{filtered}}(k-1) + (1-\alpha) \cdot damage_{debris}(k)$$

$$\alpha = \begin{cases} \max\left(bound_{lower}, \alpha \cdot scaler_{decay}\right) & \text{if } damage_{debris_{filtered}}(\mathrm{k-1}) \leq \mathrm{damge}_{debris}(k) \\ \min\left(bound_{upper}, \alpha \cdot scaler_{increase}\right) & \text{otherwise} \end{cases}$$

Typical values for the threshold and fixed quantities are
$bound_{lower} = 0.1$
$bound_{upper} = 0.99$
$scaler_{decay} = 0.99$
$scaler_{increase} = 1.02$

APPLICATION

The prognostic reasoner has been tested in experiments that model a simulated, cyclic mission profile. Figure 9 shows the assembled load and speed trajectories, which was reflective of about 40 cycles in the load-speed space, with dwells at certain set points. An indent was added to the outer race of a production bearing, which was then run under those conditions in a test rig. The bearing was examined several times during the course of the test, and actual spall length was recorded. The test ran to cage failure.

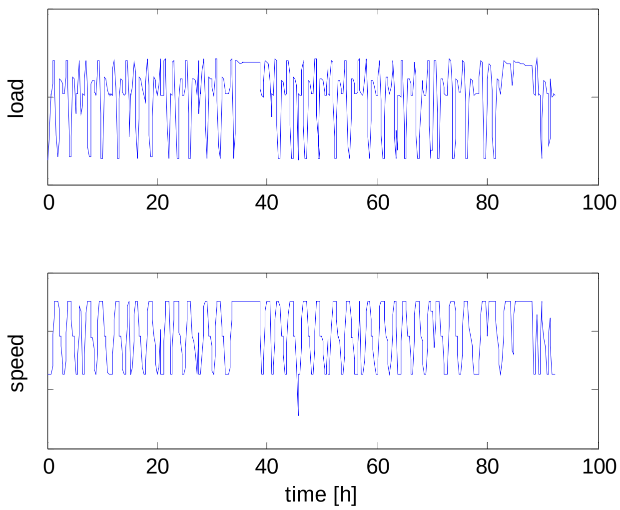


Figure 9 – Test Profile (load and speed)

As mentioned before, the fundamental characteristic of the forward confidences is that they drop as a function of time. In addition, there is an *a priori* bias assigned to the different confidences, which in turn reflects the accuracy of the models as observed during testing.

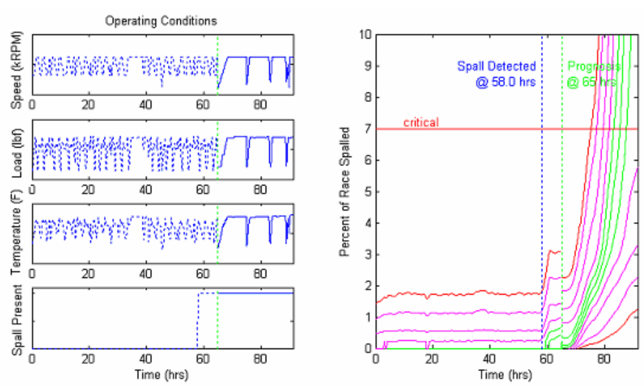


Figure 10 – Effect of future high speed and load conditions on remaining life estimates

Subplots on the left side of Figure 10 show the load, speed, and temperature conditions as recorded during the tests. The bottom subplot of Figure 10 shows the output of the diagnostic reasoner. Specifically, spall is detected at t=58.0 hours and indicated by setting the "Spall Present" flag to "1". The prognostic forward mode can be executed at any time after fault initiation. Here, we choose to execute the prognostic functions at t=65 hours. The plot on the right side of Figure 10 shows the damage estimate prior to the prognostic estimate at t=65 hours which is the output of the diagnostic reasoner. In the graph, the green lines represent the 40th, 50th, and 60th percentiles, respectively. The pink lines represent the 10th, 20th, and 30th as well as the 70th, 80th, and 90th percentiles. Finally, the red line represents the 5th and 95th percentiles. The model can output any other percentile as well, such as the percentile associated with 3σ or any other risk limit. The prognostic reasoner assesses the damage from time t=65 hours forward, using the expected load and speed profile as input (to which the uncertainty was added as described earlier). The lines from t=65 hours and greater represent the output of the prognostic reasoner. In this experiment, actual cage failure occurred at t = 93 hours. The prognostic estimate tripped the critical damage line in agreement with the experiments. A user could take action at a predetermined risk limit. In case of the 95th percentile, this would equate to about $t_{critical}$ =73 hours (i.e., where the 95th percentile crosses the critical damage line). That implies the equipment can be operated within the risk interval with these load and speed conditions for another 8 hours. Since the prognostic horizon is dependent on the future speed and load, a different speed and load profile will allow the operator to influence the remaining life of the equipment. Consider the different load and speed profile shown in Figure 11. Here, lower loads and speeds are considered for the future. The prognostic horizon increases accordingly to a larger value, implying that the equipment can be used that much longer with the same level of risk.

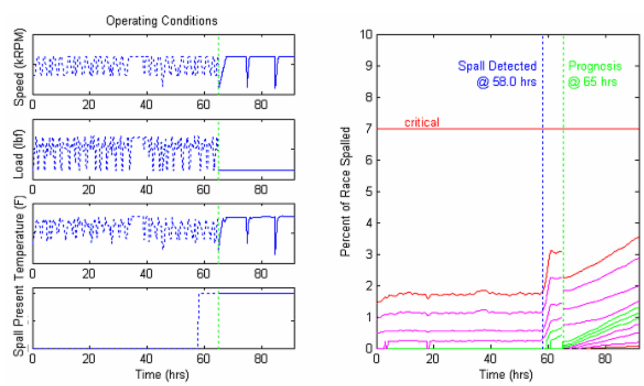


Figure 11 – Effect of future low load and speed conditions on remaining life estimates

SUMMARY & CONCLUSIONS

This paper describes how two fundamentally different methods can be aggregated to more reliably estimate remaining life and how their independent estimates can be

fused. One method uses first principles to model fault propagation through consideration of the physics of the system. The other method is an empirical model using data from experiments at known conditions and component damage level to estimate condition-based fault propagation rate. These two approaches are fused in the prognostic mode to produce a result that is more accurate and more robust than either method alone. The fusion method employs a combination of damage PDFs, subjective quality assessments, and a kernel-based regression through time. The diagnostic reasoner uses the same fusion method but adds a debris-based damage estimate and a vibration-based damage estimate to the estimation suite. The diagnostic reasoner also detects spall based on a combination of debris and vibration features. Results from rig tests where a bearing was run under mission typical flight profiles were used to validate the approach. To that end, spall was initiated and bearing spall growth was carefully monitored. Results from these tests were compared to the prognostic estimates of the reasoner and found to be in close agreement.

REFERENCES

M. Ashby and W. Scheuren, “Intelligent Maintenance Advisor for Turbine Engines (IMATE)”, Proceedings of the IEEE Aerospace Conference, 11.0309, 2000.

P. Bonanni, K. Goebel, G. Moscarino, Prognostic Reasoner for Bearings , GRC Disclosure Letter RD33952, 10/15/2004.

C. Byington, T. Merdes, and J. Kozlowski, “Fusion Techniques for Vibration and Oil Debris/Quality in Gearbox Failure Testing”, Proceedings of the International Conference on Condition Monitoring ’99, pp. 113 – 118, 1999.

P. Dempsey, R. Handschuh, and A. Afjeh, “Spiral Bevel Gear Damage Detection Using Decision Fusion Analysis”, Proceedings of the Fifth International Conference on Information Fusion, Fusion 2002, vol. 1, pp. 94-100, 2002.

Y. Freund and R. Schapire, “A Short Introduction to Boosting”, J. Japanese Society Artificial Intelligence, Vol. 14, No.5, pp. 771-780, 1999.

A. Garga, K. McClintic, R. Campbell, C.-C. Yang; M. Lebold, T. Hay, C. Byington, “Hybrid Reasoning for Prognostic Learning in CBM Systems”, IEEE Proceedings Aerospace Conference, vol. 6, 10-17 March 2001, pp. 2957 – 2969, 2001.

K. Goebel, V. Badami, M. Ashby “Information Fusion of Classifiers in Systems with Partial Redundant Information”, US patent US6757668, issued 6/23/2004.

K. Goebel, P. Bonanni, N. Eklund, “Towards an Integrated Reasoner for Bearings Prognostic”, Proceedings of 2005 IEEE Aerospace Conference, paper z11_0704, 2005.

P. Howard, and J. Reintjes, “A Straw Man for the Integration of Vibration and Oil Debris Technologies”. Helicopter Health and Usage Monitoring Systems Workshop, G. Forsyth, ed., Defense Science and Technology Organisation Geeneral Document 197, no.1, pp. 131-136, 1999.

J.-S.R. Jang, "ANFIS: Adaptive-network-based fuzzy inference systems," IEEE Transactions on Systems, Man, and Cybernetics, vol. 23, no. 3, pp. 665--685, 1993.

J. Littles and M. Buczek, “Engine System Prognosis”, Proceedings of Materials Science & Technology 2004, September 26-29, 2004, New Orleans, Louisiana, 2004.

A. Loskiewicz-Buczak, R. Uhrig, “Decision Fusion by Fuzzy Set Operations”, Proc. third IEEE Conf. Fuzzy Systems, Vol. 2, pp.1412-1417, 1994.

E. Nadaraya, “On estimating regression”. Theory of Probability and its Applications. 10, 186-190, 1964.

M. Nelson and K. Mason, “A Model-Based Approach to Information Fusion”. Proc. Information, Decision, and Control, pp. 395-400, 1999.

R. Orsagh, J. Sheldon, and C. Klenke, “Prognostics/ Diagnostics for Gas Turbine Engine Bearings”, Proceedings of ASME Turbo Expo 2003, Power for Land, Sea and Air, June 16 – 19, 2003, Atlanta, GA. GT2003-38075.

S. Petit-Renaud and T. Denoeux, “Nonparametric Regression Analysis of Uncertain and Imprecise Data Using Belief Functions”, International Journal of Approximate Reasoning, Vol. 35, No. 1, pp. 1-28, 2004.

N. S. V. Rao, “Finite Sample Performance Guarantees of Fusers for Function Estimators, Information Fusion, Vol. 1, no. 1, pp. 35-44, 2000.

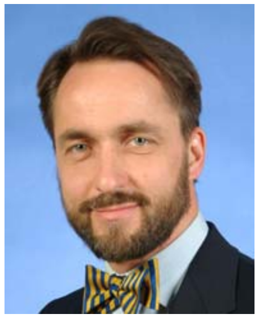
P. Smets, “What is Dempster-Shafer’s model?” Advances in the Dempster-Shafer Theory of Evidence, Yager, R., Fedrizzi, M., and Kacprzyk, J., (Eds.), John Wiley & Sons, New York, pp. 5-34, 1994.

G. Watson, "Smooth Regression Analysis", Sankhia, Series A, 26, 359 – 372, 1964.

K. Goebel, N. Eklund, and P. Bonanni, Fusing Competing Prediction Algorithms for Prognostics, Proceedings of 2006 IEEE Aerospace Conference, 11.1004, 2006.

S. Marble, Bearing Prognostics, Proceedings of 2006 IEEE Aerospace Conference, XX.YYYY, 2006.

BIOGRAPHY



Kai Goebel *received the degree of Diplom-Ingenieur from the Technische Universität München, Germany in 1990. He received the M.S. and Ph.D. from the University of California at Berkeley in 1993 and 1996, respectively.*

Dr. Goebel joined General Electric's Corporate Research and Development facility in Schenectady, NY in 1997 as a computer scientist after working as a visiting postdoctoral fellow at UC Berkeley from 1996 to 1997. He has carried out applied research in the areas of artificial intelligence, soft computing, and information fusion. His research interest lies in advancing these techniques for real time monitoring, diagnostics, and prognostics. He has fielded numerous applications for aircraft engines, transportation systems, medical systems, and manufacturing systems. He has published more than 50 technical papers in these areas.

Dr. Goebel is an adjunct professor of the CS Department at Rensselaer Polytechnic Institute (RPI), Troy, NY, since 1998 where he teaches classes in Soft Computing and Applied Intelligent Reasoning Systems.

Neil Eklund *received B.S. in 1991, two M.S. degrees in 1998, and a Ph. D. in 2002, all at the Rensselaer Polytechnic Institute.*

Dr. Eklund was a research scientist at the Lighting Research Center from 1993 to 1999. He was in the network planning department of PSINet from 1999 to 2002, before joining General Electric Global Research in Niskayuna, NY in 2002. He has worked on a wide variety of research projects, including early detection of cataract using intraocular photoluminescence, multiobjective bond portfolio optimization, and on-wing fault detection and accommodation in gas turbine aircraft engines. His current research interests involve developing hybrid soft/hard computing approaches for real-world problems, particularly real time monitoring, diagnostics, and prognostics.

Dr. Eklund is an adjunct professor in the Engineering/CS department of the Graduate College of Union University, Schenectady, NY, since 2005 where he teaches classes in Computational Intelligence and Machine Learning.

Pierino Bonanni *received his S.B., S.M., E.E., and Ph.D. degrees in Electrical Engineering from the Massachusetts Institute of Technology in 1983, 1983, 1984, and 1991, respectively.*

Dr. Bonanni joined GE Global Research Center in Niskayuna, New York, in 1991. His specialization is in signal and image processing, with emphasis on detection, estimation, and inverse problems arising in controls and diagnostics applications. He has performed and led a variety of applied research efforts in fault detection. machine vision, robotics, automated manufacturing, satellite navigation and control, non-destructive inspection, and remote sensing. He is the recipient of several GE team and project management awards, and holds 7 U.S. patents.